Report #2
PD-Internship
Lerner Paul
18/03/2019

I mostly read papers about PD and Handwriting (HaW) for now. I highlighted the parts where something is unclear or I doubt and **need feedback**…
We also created a GitLab to share the code and the reports[1]. I managed to load and explore the data from PaHaW.

Summary

[1] https://gitlab.telecom-paristech.fr/laurence.likforman/parkinson_detection

Lexicon

Sensitivity (Se) refers to the test's ability to correctly detect ill patients who do have the condition (Altman & Bland).
Specificity (Sp) relates to the test's ability to correctly reject healthy patients without a condition.
If we define :

- True positive (TP) as sick people correctly identified as sick
- False positive (FP) as healthy people incorrectly identified as sick
- True negative (TN) as healthy people correctly identified as healthy
- False negative (FN) sick people incorrectly identified as healthy

then :

$$Se = \frac{TP}{TP + FN} \quad ; \quad Sp = \frac{TN}{TN + FP}$$

Datasets

A summary of the different datasets characteristics is available here and in the git repo.

Handwriting hallmarks of PD

Micrographia

Graça et al. cite Chauduri et al. 2006 to report that *"micrography [...] may not be truly related to PD"* but, reading the paper from Chauduri et al., I wasn't able to find anything that corroborates this statement.

Inzelberg et al. report that micrographia is not exclusive to PD and has often been observed by different researchers (cf. article for references) following focal basal ganglia lesions without any accompanying parkinsonism. They made this figure which resumes all the disease which cause micrographia :

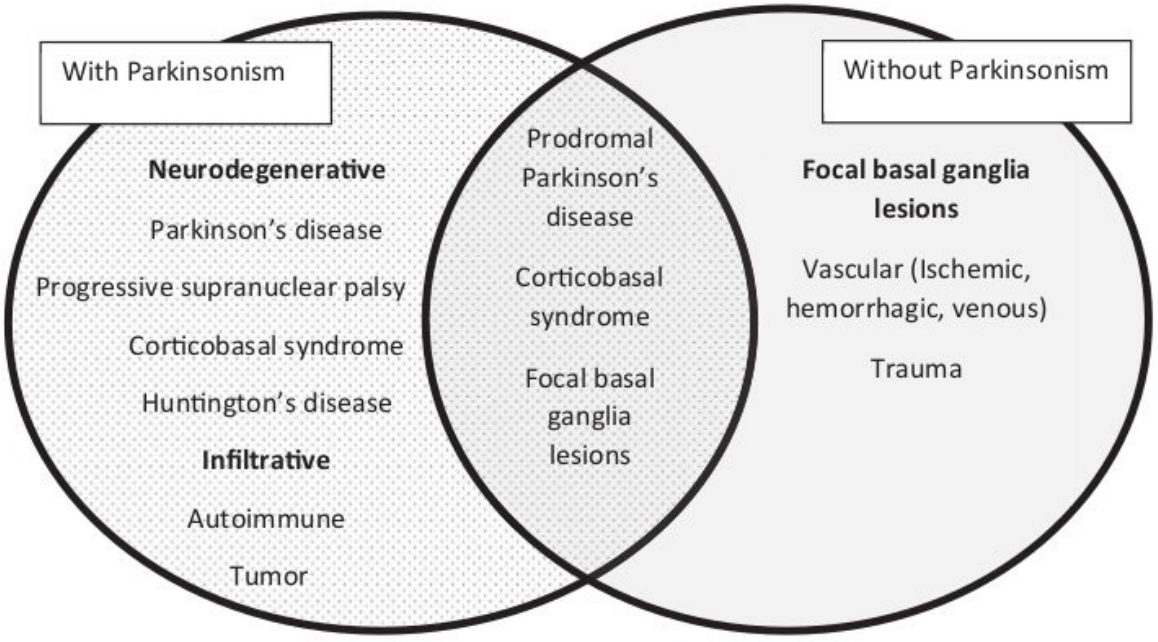

Fig. 2. Etiology of micrographia with and without accompanying Parkinsonism.

McLennan et al. found that micrographia may antedate additional motor signs of PD by three to four years. They also observed that 30% of patients showed micrographia in the course of PD and about 5% reported its occurrence prior to PD diagnosis. Therefore, one should be careful about collecting data from patients who specifically suffer from micrographia (PaHaW and PDMultiMC datasets don't mention anything about their subject suffering from micrographia). Although the number of patients showing micrographia varies according to the study's methodology, such as information retrieved through verbal history (about 60% of patients), actual testing (about 50%) (Shukla et al.).
Wu et al. show that asking patients to **pay attention to the letter size** improves **micrographia**. In the same way, Berardelli et al. show that attention to movement is beneficial to **bradykinesia**. Therefore, one should be careful to not give any hint to subjects when collecting data. Drotar et al. write that *"subjects were asked to write a sentence at a **self-determined comfortable size and speed**."*. Kotsavasiloglou et al. subjects were instructed to keep the pen's **velocity as constant as possible** (they justify this because *"The effects of PD are dominant particularly at rest and during steady movements"*) .

Dysgraphia

Moetesum et al. 2018 define **bradykinesia** as :

"Bradykinesia or slowness of movement (either due to motor or cognitive dysfunction), causes a potential PD patient to complete a graphomotor task in more time than usually required."
And **tremors** as :
"involuntary to and fro movements that can be visualized by irregular formations of characters and drawings."
Despite this, Lettaneux et al. 2014 don't include bradykinesia and tremor in **dysgraphia**.

I wonder if dysgraphia has the same properties (cf. Inzelberg et al.) as micrographia :

- 2 types : progressive & consistent
- affected by writing directions (i.e. horizontal or vertical)
- antedate additional motor signs of PD
- affected by dominant hand
- affected by language (i.e. native or learned)

San Luciano et al. 2016 show that over 138 patients, only 32% had predominant symptoms on their **dominant** side vs. 34% in their **non-dominant** side and 34% had symmetric disease. This is good news since all PaHaW subjects used only their **dominant right-hand**.
Letanneux et al. show that writing velocity and smoothness/fluency abnormalities are more frequent (above 75%) than diminished letter size (between 30%-50%), this is an excellent motivation to analyse kinematic features. They also conclude that *"dopamine depletion would affect not only the amplitude of the motor output, but also its kinematics"*.
Dysgraphia is promising to discriminate between PD and other diseases as Yu et al. reported significant differences in handwriting kinematics between patients with PD and Essential Tremor (ET) ; Ling et al. reported difference between PD and Progressive Supranuclear Palsy (PSP) ; PD patients react to Levodopa unlike non-PD patients so I guess one could collect data from people suffering from different disease both on-drug and off-drug… Although Inzelberg et al. report that L-dopa effects are inconsistent and clinical observations states that only some PD patients restore letter size with dopaminergic therapy (cf. article for references). Also, Wu et al. show that L-dopa improves consistent micrographia but not progressive micrographia.

HaW Features for PD Diagnosis

Graça et al. 2014 had interesting results by collecting multimodal data through a smartphone, they studied :

- tremor, rigidity and bradykinesia through Archimedean spiral
- rigidity and bradykinesia through a tap game
- freezing of gait through a walking exercise (used accelerometer measures from the smartphone)

they achieved 90% sensitivity (10 fold-cross validation) with Classification Rules (RipperK). Even though their dataset is very small (17 PD and 18 controls) these results are promising.

Their goal is to provide a simple method to be used by General Practitioners that may confirm the possible **suspicion** of PD and, in this way, sending the patient to a Neurologist as early as possible, thus allowing the beginning of treatment in earlier phases of the disease. Unfortunately, since 2014, the authors didn't further published about PD diagnosis.

San Luciano et al. 2016 collected data from 138 PD (including 50 early stage) and 150 controls. The one and only task was Archimedean spiral from both hands. They used features from the difference between the 2 hands. 46% PD patients were on medication, the rest off. They chose to include on-medication patients in order to *"stress the potential of this technique to discern between milder forms of PD and controls, as treated PD subjects may closely resemble controls."*. They achieved 86% sensitivity (Se), 81% specificity (Sp) and 84% area under the curve (AUC).
Moreover, they also tried to classified only **early stage PD** (less than 4 years of disease duration) and obtained 81% AUC.
Also, they performed a sensitivity analysis limited to **men alone** and found the same direction and significance of results. This is a good news for the PDMultiMC dataset which has an unequal number of men among PD and controls.

Taleb et al. 2017 averaged the training data of each segment (i.e. word) of the same task for each subject. On March 7th we agreed upon keeping each segment separately in order to augment the dataset and not lose any information. Thus we will have to decide how to merge the model prediction over each segment to classify a given subject. Although I'm not sure if the data is already splitted into segment, we could also simply keep the whole task as a training example.

Drotar et al., 2014 found that kinematic features from in-air movements were more discriminative than features from on-surface movements (87% VS 78% sensitivity with a SVM classifier). This is explained by Rosenblum et al. 2013 as *"in-air time is a manifestation of* 'planning the next movement'*, as required in the sequential process of handwriting"*.
Drotar et al., 2015 found that their second task was *"more contributing to overall classification performance"* and they explain that because : *"The second task is a pseudoword not existing in Czech language, whereas tasks from 3 to 6 include words used in everyday language. Therefore, the performance of the second task* ***is not automatized****. The lateral premotor cortex is* ***additionally activated*** *when attention must be paid during a motor task."*.
This goes against everything else I've been reading, if the task is not automatized it should be **the less discriminative**. Letanneux et al. recommend *"Asking participants to write a long text should be* ***avoided*** *because it implies either copying the text or writing under dictation, and in both cases,* ***cognitive processes are required****. Participants should write single* ***familiar words****, without any spelling difficulty and syntactic and semantic contexts."*. Moreover, I doubt that this difference between task 2 and 3 is really meaningful given that task 2 is to write "le" and

task 3 is to write "les"... Moreover, Moetesum et al. 2018 found, on the contrary, that the "les" task was more discriminative than the "le" task (60% VS 57% accuracy).
Drotar et al., 2016 found that the Archimedean spiral task was the least discriminative task although Stanley et al. suggests that it might be an early marker for PD. Drotar et al. explain that it may be because they focused on handwriting structures and didn't use any **spiral-specific** features. On the contrary, Moetesum et al. 2018 found that the Archimedean Spiral task was the most discriminative. Unlike Drotar et al. they used a CNN to extract the features. The advantage of the spiral task over handwriting ones is that it is more universal, not influenced by language or education (Saunders-Pullman et al.).
Drotar et al., 2016 also found that pressure features were more discriminative than kinematic features (82.5% VS 75.4% accuracy, cf. Report #1), it would be interesting to see if we get the same results with a neural network : Pereira et al. don't report experiments separating the measures ; Mucha et al. 2018 only used kinematic features and Taleb et al. don't report which of the features are the most discriminative.

Kotsavasiloglou et al. 2016 achieved 88% sensitivity (Se) and 95% specificity (Sp) on a 44 subject dataset where the tasks were to draw simple horizontal lines. They partly chose this task because it *"allows for both hands to be used and assessed, constituting the dexterity of the dominant hand irrelevant to the assessment"*. The most discriminative features are the normalized velocity variability (a new metric that they introduced) and the entropies of the horizontal and vertical components of the signal. They didn't use in-air movements (not compatible with their task) nor pressure, unlike Drotar et al.

Mucha et al. 2018 achieved 89 % sensitivity, and 91 % specificity on a modified **PaHaW** dataset of 33 PD and 36 controls with an extra task : "repetitive loops" (though this task didn't provide the best classification results). The authors don't even mention that this dataset is not exactly the same as the original PaHaW… Their results are consistent with those of Drotar et al. in the sense that they found that the sentence task was the most discriminative task. They outperformed the results of Drotar et al., even though they didn't use pressure features, thus demonstrating the power of fractional order derivation (FDE, Podlubny et al.).

Pereira et al. 2018 recolted data for the **NewHandPD** dataset :

1) spirals
2) meanders
3) circle in the air
4) circle on the paper
5) right-handed diadochokinesis
6) left-handed diadochokinesis

From these 6 tasks they found that the most discriminatives were spirals and meanders (cf. paper for classification results). Is it possible to treat every data as a training example, regardless of the task ? (Pereira/Passos/Moetesum et al. don't report trying that).

Classification

Summary of the different results

The best results for each single dataset are printed in **bold**.

Dataset	Authors	Classifier	Se (%)	Sp (%)	Evaluation method
San Luciano	San Luciano et al. 2016	Linear mixed effect	**86**	**81**	3-fold cross validation with 3 runnings
PDMultiMC	Taleb et al. 2017	SVM	**94**	**100**	4-fold cross validation
Kotsavasiloglou	Kotsavasiloglou et al. 2016	Naïve Bayes	**88**	**95**	10-fold cross validation
PaHaW	Drotar et al. 2016	SVM	87	81	10-fold cross validation
PaHaW	Moetesum et al. 2018	CNN to extract features then majority voting SVM	84	82	10-fold cross validation
Modified PaHaW	Mucha et al. 2018	FDE to compute features then random forests classifier	**89**	**91**	7-fold cross-validation
HandPD	Pereira et al. 2016, October	CNN	91	76	random 75-25 split with 20 runnings

HandPD	Passos et al. 2018	CNN to extract features then OPF	90	**97**	five-fold cross-validation
"HandPD2"	Pereira et al. 2016, December	CNN	91	89	random 50-50 split with 10 runnings
NewHandPD	Pereira et al. 2018	majority voting CNN	**97**	94	?

Statistical analysis

Rosenblum et al. 2013 achieved 95% sensitivity and 100% specificity using discriminant analysis (can we compare those results and the rest of literature ? e.g. no train-test set, no cross valid). Drotar et al. 2014, 2015 and Mucha et al. 2018 cite this paper only as a *"preliminary data [study that] suggest that handwriting might serve as a diagnostic marker for PD"* even though they end up with worse results than them. Unlike them, Moetesum et al. provide the results of Rosenblum et al. in a comparative way even though they mention that *"meaningful comparison"* of their system is only possible with the works of Drotar et al. Rosenblum et al. 2013 used only 9 features : mean time taken to write each stroke, both on-paper and in-air ; stroke time, both on-paper and in-air ; mean velocity for the entire task in seconds ; mean stroke's length, width, and height in centimeters ; and the mean pressure.
The 40 subjects did only two tasks : write their name and copy an address (same address for all). Patients performed study while on medication.

Manually-extracted Feature Machine Learning

Please refer to Summary of the different results.

Kotsavasiloglou et al. 2016 achieved better results with Bayes than with AdaBoost and SVM, unlike Drotar et al. Though I'm not sure that we can make a direct comparison of the algorithms given that they were not applied on the same dataset.

Even though their Mucha et al. are inferior to Taleb et al. 2017 they say they can't make a *"relevant comparison"* because they have a significantly greater number of samples.

Taleb et al. 2018 achieved 94%, 92%, and 88% accuracy predicting the Hoehn and Yahr (H&Y) stage, UPDRS scores, and total UPDRS, respectively. This motivates the idea that HaW analysis can be used to monitor PD progression in a cheap, non invasive and possibly "from home" way. Although, further studies with larger datasets should be used to confirm this. The PDMultiMC is an equally balanced dataset of 32 subjects, thus significantly smaller than **PaHaW**, moreover, there are 7 stages in the modified H&Y stage (with stage 1.5 and 2.5, excluding the 0 since Taleb et al. focused on the PD only) and PDMultiMC contains only 5 of them : none of the PD has a 2.5 nor a 5 H&Y stage. In the same way, most of the possible UPDRS scores are not represented in the PDMultiMC dataset.

On March 7th we decided that we will try a pure data-driven approach were we **don't extract any features** from the data, as we're going to use deep learning techniques, unlike most of the literature. Although we're a priori going to normalize the data before feeding it to the model, like Pereira et al. and Drotar et al. Maybe we could apply a Fourier transform also ?

CNN-extracted Feature Machine Learning

Passos et al. 2018 used a CNN : ResNet-50 (trained for object detection on ImageNet) to extract features from **the images of the HandPD** dataset which were then fed to a classifier (notice that they use **HandPD** and not **NewHandPD**, one can assume that NewHandPD was not available at this time). The authors experimented with OPF, SVM and Bayes classifiers. It's a bit unclear what they call "patient accuracy" but I think it's the precision (aka Positive predictive value, PPV). In the same way, I think "control accuracy" is the Negative Predictive Value (NPV).
I'm very surprised that they obtain 84% Se and 92% Sp using only Res-Net (i.e. by adding just one classification layer after the feature extraction) because the data and the classification task are very different from ImageNet (Pereira et al. chose not to use **transfer learning** because of that). They achieved 89% Se 97% Sp with OPF after applying a PCA to reduce the extracted features to a dimensional space of size 100. This is the best Sp on HandPD/NewHandPD to this day. Their better results are on the spiral task but their results on the meander task are quite similar. Their results are more robust than the ones of Pereira et al. 2016 : they deal better with the imbalanced data. Moreover, they achieved a great Sp even though there is a very little number of controls in the HandPD dataset (which caused the inconsistent results of Pereira et al. 2016). It's also impressing that they used only the **images of the HandPD** dataset, and not the sensors data, Pereira et al. 2015 had had very poor results doing that.

In the same way, Moetesum et al. used a CNN, AlexNet (also trained on ImageNet) to extract features from the **PaHaW** dataset before feeding them to a SVM. Even though they use PaHaW, they don't use the sensors measures but only the trace (generated thanks to the

recorded positions of the pen). This is because their work aims at using only the static features (in opposition to dynamic features like velocity, etc.). Thus the examination could be done using only a regular pen and pencil (even though the tablets and smart pen are already acknowledged as "cheap" or "inexpensive"). Like Pereira et al. 2018 they train a model over each task then combine the prediction of each model using **majority voting**.
Their work is quite similar to Passos et al. but they don't cite each other, the papers both came out in 2018 so one can assume that they were not aware of each other's works.

Feature-Free Machine Learning

Pereira et al. 2016, October attained 91% sensitivity and 76% specificity, using convolutional neural networks (CNN) by transforming exam signals into a grayscale image (1 row per milliseconds and 6 columns corresponding to the 6 signal channels). They also had very strange results, e.g. 98% sensitivity but 6% specificity with some architectures. This is probably because their dataset is imbalanced. They trained their model on only 2 tasks for now : spiral and meander (their best results are on the latter). It's interesting that they achieved decent results with an Optimum-Path Forest (OPF, Papa et al. 2009) over the raw data : 87% sensitivity and 71% specificity. These results serve as a baseline. Also, their results with OPF are more consistent/robust : neither sensitivity nor specificity gets worse than 50% depending on the task, image size and dataset split ratio, unlike their different CNN architectures.
The paper came out in 2016 and they claim to be the first to diagnose PD using **deep learning** techniques. They chose not to use **transfer learning** techniques since their images are "domain-specific" (they don't mention the PaHaW dataset nor any work from Drotar et al. in this paper but they do in their latest work and still chose not to use transfer learning for the same reasons).
In a more recent work (Pereira et al. 2016, December) they achieved 91% sensitivity and 89% specificity by fine-tuning the hyperparameters : learning rate η, penalty parameter (momentum) α and weight decay λ using the Bat Algorithm (BA), Particle Swarm Optimization (PSO) and Firefly Algorithm (FA). Their best result were achieved with BA. This time, their best results are with the spiral task, and not the meander. Their better results compared to their previous work might be explained because of their use of an updated version of the **HandPD** dataset (let's call it **HandPD2**) which is more balanced than the old one.
In their latest work (Pereira et al. 2018) they achieved 97% sensitivity and 94% specificity by taking into account more tasks (i.e. they train a CNN over each task) and predicting by combining the prediction of each CNN. Combination is achieved through **majority voting**. This is the best Se on HandPD/NewHandPD to this day. They say that the CNN parameters were set empirically and strangely don't mention their previous work on CNN finetuning (one can assume that they used the parameters that they found best in this previous work). As a

baseline, they also used OPF, SVM and Naive Bayes classifier using the same voting technique but transforming the data with gray level co-occurrence matrices (GLCM) : *"GLCM stands for the distribution of co-occurring pixels values concerning a given offset. For our purpose, we used the* ***energy, entropy, contrast, homogeneity and correlation*** *features computed over the matrices built upon the angles as 0° , 45°, 90° and 135°"*. They also achieved great results with those : 99% sensitivity and 91% specificity. Once again, their better results compared to their previous work might be explained because of their use of an updated version of the HandPD dataset : **NewHandPD** which is more balanced than **HandPD** and contains more observations than **HandPD2**.

Our Goal - Conclusion

We should state clearly that our goal is to classify healthy and sick people and not people suffering from PD and people suffering from other diseases which cause micrographia (listed in Fig. 2), bradykinesia or tremor as we don't have any dataset containing those (though we should check if it exists and, dysgraphia is promising to discriminate between PD and other diseases as explained in the Dysgraphia section). Therefore, our work is an extension of the work of Drotar et al. and Pereira et al. and our goal is to find a better machine learning tool than them. In order to make our work more interesting I think we should really try cross learning between PaHaW, PDMultiMC and NewHandPD. **Or** cross learning between medicated and drug-free patients. I think I should chat with Catherine and maybe work with her depending on her goals.

Even though it's obvious , we should maybe state that our work in entirely focused on **motor symptoms** of PD, as *"Non-motor symptoms correlate with advancing age and disease severity, although some non-motor symptoms, such as olfactory problems, constipation, depression, and rapid eye movement disorder, can occur* ***early in the disease*** *(panel 1). As the average age and life expectancy of the population increases,* ***the non-motor features of Parkinson's disease become increasingly important***." (Chauduri et al. 2006)

I'll have to read about deep learning and recurrent neural networks but I can already see several ways of representing the data / training the model :

- vector containing only one measure (e.g. pressure)
- vector containing all the measures (e.g. x, y, pressure, etc.)
- transform the measures by applying a fourier transform
- train the model over each task then merge the prediction. How ?
    - average or max pooling over the hidden states
    - majority voting
- use all tasks to train the model at the same time :

- treat every task as a training example (unlikely to work)
- concatenate all the tasks in one vector (row-wise or column-wise or "measure"-wise ?)

Obviously this is just food for thought and we should discuss and investigate research before trying any of that.

Todo List

On March 7th and 12th we agreed that it would be interesting that I get into papers about :

- automatic diagnosis of other diseases based on HaW features
- automatic diagnosis of PD using other features than HaW, e.g. speech
- automatic diagnosis (via HaW ?) of non-PD disease that cause micrographia (listed in Fig. 2).

I think I should ask to San Luciano et al. if they're willing to share their dataset, as it's much bigger than PaHaW.

I also think that I should look into the conferences where the literature published their papers so maybe we could attend to one.

References

Berardelli, A., Rothwell, J. C., Thompson, P. D., & Hallett, M. (2001). Pathophysiology of bradykinesia in Parkinson's disease. *Brain*, *124*(11), 2131-2146.

Chaudhuri, K. R., Healy, D. G., & Schapira, A. H. (2006). Non-motor symptoms of Parkinson's disease: diagnosis and management. The Lancet Neurology, 5(3), 235-245.

Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2014). Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease. *Computer methods and programs in biomedicine*, *117*(3), 405-411.

Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2015). Decision support framework for Parkinson's disease based on novel handwriting markers. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*(3), 508-516.

Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2016). Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artificial intelligence in Medicine*, *67*, 39-46.

Graça, R., e Castro, R. S., & Cevada, J. (2014, June). Parkdetect: Early diagnosing parkinson's disease. In *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (pp. 1-6). IEEE.

Haralick, R. M., & Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.

Inzelberg, R., Plotnik, M., Harpaz, N. K., & Flash, T. (2016). Micrographia, much beyond the writer's hand. *Parkinsonism & related disorders*, *26*, 1-9.

Kotsavasiloglou, C., Kostikis, N., Hristu-Varsakelis, D., & Arnaoutoglou, M. (2017). Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomedical Signal Processing and Control*, *31*, 174-180.

Letanneux, A., Danna, J., Velay, J. L., Viallet, F., & Pinto, S. (2014). From micrographia to Parkinson's disease dysgraphia. *Movement Disorders*, *29*(12), 1467-1475.

Ling, H., Massey, L. A., Lees, A. J., Brown, P., & Day, B. L. (2012). Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain*, *135*(4), 1141-1153.

McLennan, J. E., Nakano, K., Tyler, H. R., & Schwab, R. S. (1972). Micrographia in Parkinson's disease. *Journal of the neurological sciences*, *15*(2), 141-152.

Moetesum, M., Siddiqi, I., Vincent, N., & Cloppet, F. (2018). Assessing visual attributes of handwriting for prediction of neurological disorders—A case study on Parkinson's disease. *Pattern Recognition Letters*.

Mucha, J., Mekyska, J., Faundez-Zanuy, M., Lopez-De-Ipina, K., Zvoncak, V., Galaz, Z., ... & Rektorova, I. (2018, November). Advanced Parkinson's Disease Dysgraphia Analysis Based on Fractional Derivatives of Online Handwriting. In *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* (pp. 1-6). IEEE.

Papa, J. P., Falcao, A. X., & Suzuki, C. T. (2009). Supervised pattern classification based on optimum-path forest. *International Journal of Imaging Systems and Technology*, *19*(2), 120-131.

Passos, L. A., Pereira, C. R., Rezende, E. R., Carvalho, T. J., Weber, S. A., Hook, C., & Papa, J. P. (2018, May). Parkinson disease identification using residual networks and optimum-path forest. In *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)* (pp. 000325-000330). IEEE.

Pereira, C. R., Pereira, D. R., da Silva, F. A., Hook, C., Weber, S. A., Pereira, L. A., & Papa, J. P. (2015, June). A step towards the automated diagnosis of parkinson's disease: Analyzing handwriting movements. In *2015 IEEE 28th International Symposium on Computer-Based Medical Systems* (pp. 171-176). IEEE.

Pereira, C. R., Weber, S. A., Hook, C., Rosa, G. H., & Papa, J. P. (2016, October). Deep learning-aided parkinson's disease diagnosis from handwritten dynamics. In *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (pp. 340-346). Ieee.

Pereira, C. R., Pereira, D. R., Papa, J. P., Rosa, G. H., & Yang, X. S. (2016, December). Convolutional neural networks applied for parkinson's disease identification. In *Machine Learning for Health Informatics* (pp. 377-390). Springer, Cham.

Pereira, C. R., Pereira, D. R., Rosa, G. H., Albuquerque, V. H., Weber, S. A., Hook, C., & Papa, J. P. (2018). Handwritten dynamics assessment through convolutional neural networks: An application to Parkinson's disease identification. *Artificial intelligence in medicine*, *87*, 67-77.

Podlubny, I. (1998). *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications* (Vol. 198). Elsevier.

Rosenblum, S., Samuel, M., Zlotnik, S., Erikh, I., & Schlesinger, I. (2013). Handwriting as an objective tool for Parkinson's disease diagnosis. *Journal of neurology*, *260*(9), 2357-2361.

San Luciano, M., Wang, C., Ortega, R. A., Yu, Q., Boschung, S., Soto-Valencia, J., ... & Saunders-Pullman, R. (2016). Digitized spiral drawing: A possible biomarker for early Parkinson's disease. *PloS one*, *11*(10), e0162799.

Saunders-Pullman, R., Derby, C., Stanley, K., Floyd, A., Bressman, S., Lipton, R. B., ... & Pullman, S. L. (2008). Validity of spiral analysis in early Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society*, *23*(4), 531-537.

Shukla, A. W., Ounpraseuth, S., Okun, M. S., Gray, V., Schwankhaus, J., & Metzer, W. S. (2012). Micrographia and related deficits in Parkinson's disease: a cross-sectional study. *BMJ open*, *2*(3), e000628.

Stanley, K., Hagenah, J., Brüggemann, N., Reetz, K., Severt, L., Klein, C., ... & Saunders-Pullman, R. (2010). Digitized spiral analysis is a promising early motor marker for Parkinson Disease. *Parkinsonism & related disorders*, *16*(3), 233-234.

Taleb, C., Khachab, M., Mokbel, C., & Likforman-Sulem, L. (2017, April). Feature selection for an improved Parkinson's disease identification based on handwriting. In *2017 1st International Workshop on Arabic Script Analysis and Recognition (ASAR)* (pp. 52-56). IEEE.

Taleb, C., Khachab, M., Mokbel, C., & Likforman-Sulem, L. (2018, March). A Reliable Method to Predict Parkinson's Disease Stage and Progression based on Handwriting and Re-sampling Approaches. In *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)* (pp. 7-12). IEEE.

Wu, T., Zhang, J., Hallett, M., Feng, T., Hou, Y., & Chan, P. (2015). Neural correlates underlying micrographia in Parkinson's disease. *Brain*, *139*(1), 144-160.

Yu, N. Y., Van Gemmert, A. W., & Chang, S. H. (2017). Characterization of graphomotor functions in individuals with Parkinson's disease and essential tremor. *Behavior research methods*, *49*(3), 913-922.